

IS6110, an IS-like element of *Mycobacterium tuberculosis* complex

D. Thierry, M.D.Cave², K.D.Eisenach², J.T.Crawford², J.H.Bates², B.Gicquel¹ and J.L.Guesdon

Laboratoire des Sondes Froides and ¹Unité de Génie Microbiologique, Institut Pasteur, 75724 Paris Cedex 15, France and ²McClellan Memorial Veterans Hospital and University of Arkansas for Medical Sciences, Little Rock, AR, USA

Submitted December 6, 1989

GenBank accession no. M29899

Mobile genetic elements are useful genetic tools. They have been found in most organisms which have been examined (for recent reviews see 1). *IS900* was isolated from *M. paratuberculosis* (2) and *IS6100* from a *M. fortuitum* strain (our unpublished results). IS elements have been used as taxonomic markers useful for diagnostic purposes, (2, 3). From a *M. tuberculosis* cosmid library constructed in pHC79 (4), an IS-like element, *IS6110*, was identified as a repeated sequence, by screening the library with labelled *M. tuberculosis* total DNA. This sequence (1361 nt) possesses characteristics of IS elements, i.e., inverted (28bp with 3 mismatched bp) and direct (3bp) repeats of the target sequence at its extremities. A search in the EMBL data bank has revealed homologies with *IS3411*, an insertion element from *E. coli* (5). Cross-hybridization was observed between *IS6110* and a repeated sequence previously isolated from *M. tuberculosis* (6). *IS6110* has been found in *M. tuberculosis* and *M. bovis* but not in any of the other mycobacteria tested (our unpublished data). Therefore, *IS6110* will be used as probe for the identification of the *M. tuberculosis* complex.

DR INV.REP.

```

CGATGAACCG CCCC GGCCATG TCCGGAGACT CCAGTTCTTG GAAAGGATGG GGT CAT GTCA GGTGGTTCAT
CGAGGAGGTA CCCGCCGGAG CTGCGTGAGC GGGCGGTGCG GATGGTCGCA GAGATCCGCG GTCAGCACGA
TTCGGAGTGG GCAGCGATCA GTGAGGTCGC CCGTCTACTT GGTGTTGGCT GCGCGGAGAC GGTGCGTAAG
TGGGTGCGCC AGGCCAGGT CGATGCCGGC GCACGGCCCG GGACCACGAC CGAAGAACATCC GCTGAGCTGA
AGCGCTTAGC GGCAGGACAA CGCCGAATTG CGAAGGGCGA ACGCGATTTC AAAGACCGCG TCGGCTTTCT
TCGCGGCCGA GCTCGACCGG CCAGCACGCT ATTAAACGGT TCATCGCCGA TCATCAGGGC CACCGCGAGG
GCCCGATGG TTTGCGGTGG GGTGTCGAGT CGATCTGCAC ACAGCTGACC GAGCTGGGTG TGCCGATCGC
CCCATCGACC TACTACGACC ACATCAACCG GGAGCCCAGC CGCCGCGAGC TGCGCGATGG CGAACTCAAG
GAGCACATCA GCCGCGTCCA CGCCGCCAAC TACGGTGTGTT ACGGTGCCCG CAAAGTGTGG CTAACCCCTGA
ACCGTCAGGG CATCGAGGTG GCCAGATGCA CCGTGAACG GCTGATGACC AAACCTGGCC TGTCCGGGAC
CACCCGCGGC AAAGCCCGCA GGACCACGAT CGCTGATCCG CCCACAGCCC GTCCCGCCGA TCTCGTCCAG
CGCCGCTTCG GACCACCGC ACCTAACCGG CTGTGGGTAG CAGACCTCAC CTATGTGTCG ACCTGGGAG
GGTTCGCCTA CGTGGCCTT GTCAACGACG CCTACGCTCG CAGGATCCTG GGCTGGCGGG TCGCTTCCAC
GATGGCCACC TCCATGGTCC TCGACGCGAT CGAGCAAGCC ATCTGGACCC GCCAACAAAGA AGGCGTACTC
GACCTGAAAG ACGTTATCCA CCATACGGAT AGGGGATCTC AGTACACATC GATCCGGTTC AGCGAGCGGC
TCGCGAGGC AGGCATCCAA CGCTCGGTGAG GAGCGGTGCG AAGCTCCTAT GACAATGCAC TAGCCGAGAC
GATCAACGGC CTATACAAGA CGGAGCTGAT CAAACCCGGC AAGCCCTGGC GGTCCATCGA GGATGTGAG
TTGGCCACCG CGCGCTGGGT CGACTGGTTC AACCATCGCC GCCTCTACCA GTACTGCGGC GACGTCCCCG
CGGTCGAACCT CGAGGCTGCC TACTACGCTC AACGCCAGAG ACCAGCCGCC GGCTGAGGTC TCAGATCAGA
GAGCTCCGG ACTCACCGGG GCGGTTACCGA

```

INV.REP.

DR

Figure 1. Complete nucleotide sequence of *IS6110*. Direct (DR) and inverted (*INV.REP.*) repeats are shown by lines.